

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 8, 2001, 13:04:36 ; Search time 5356.8 Seconds

(without alignments)
17812.975 Million cell updates/sec

Title: US-09-227-881-3
Perfect score: 6169
Sequence: 1 atcttttgttcagtttaccc... cttagtgccctccatgtcag 6169
Scoring table: OLIGO_NUC
capop 60.0 , Gapext 60.0

Searched: 1344157 seqs, 7733874588 residues
Word size : 8

Total number of hits satisfying chosen parameters: 2551636

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl,*
 1: gb_bal: *
 2: gb_ba2: *
 3: gb_ba3: *
 4: gb_bnl: *
 5: gb_bn2: *
 6: gb_bn3: *
 7: gb_cm: *
 8: gb_cv: *
 9: gb_pat1: *
 10: gb_pat2: *
 11: gb_ph: *
 12: gb_p11: *
 13: gb_p12: *
 14: gb_p13: *
 15: gb_p14: *
 16: em_ba1: *
 17: em_ba2: *
 18: em_fun: *
 19: em_htgo_hum: *
 20: em_htgo_inv: *
 21: em_htgo_rod: *
 22: em_htg_hum1: *
 23: em_htg_hum2: *
 24: em_htg_hum3: *
 25: em_htg_hum4: *
 26: em_htg_other: *
 27: em_htg_hum5: *
 28: em_htg_hum7: *
 29: em_htg_hum8: *
 30: em_htg_inv1: *
 31: em_htg_inv2: *
 32: em_htg_inv3: *
 33: em_htg_inv4: *
 34: em_htg_inv5: *
 35: em_htg_inv6: *
 36: em_htg_inv7: *
 37: em_htg_inv8: *
 38: em_hum5: *
 39: em_hum6: *
 40: em_hum7: *
 41: em_in1: *
 42: em_on: *
 43: em_or: *

44: em_cv: *
 45: em_pat: *
 46: em_ph: *
 47: em_pl: *
 48: em_ro: *
 49: em_sts: *
 50: em_sy: *
 51: em_un: *
 52: em_v1: *
 53: gb_sts1: *
 54: gb_sts2: *
 55: gb_sts3: *
 56: gb_sy: *
 57: gb_un: *
 58: gb_v1: *
 59: gb_v12: *
 60: gb_htg1: *
 61: gb_htg2: *
 62: gb_htg3: *
 63: gb_htg4: *
 64: gb_htg5: *
 65: gb_htg6: *
 66: gb_htg7: *
 67: gb_htg8: *
 68: gb_htg9: *
 69: gb_htg10: *
 70: gb_htg11: *
 71: gb_htg12: *
 72: gb_htg13: *
 73: gb_htg14: *
 74: gb_htg15: *
 75: gb_htg16: *
 76: gb_htg17: *
 77: gb_htg18: *
 78: gb_htg19: *
 79: gb_htg20: *
 80: gb_htg21: *
 81: gb_htg22: *
 82: gb_htg23: *
 83: gb_htg24: *
 84: gb_htg25: *
 85: gb_pr1: *
 86: gb_pr2: *
 87: gb_pr3: *
 88: gb_pr4: *
 89: gb_pr5: *
 90: gb_pr6: *
 91: gb_pr7: *
 92: gb_pr8: *
 93: gb_pr9: *
 94: gb_rol: *
 95: gb_rol2: *
 96: gb_in4: *
 97: gb_Prl0: *
 98: em_ba3: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
c 1	5552	90.0	79376 92	HS454c6 Z98750 Human DNA s
c 2	5300	85.9	5300 88	AF007562 Homo sapi AF007562 Homo sapi AC024490 Homo sapi
c 3	5243	85.0	170425 68	AF024490 Homo sapi AF024490 Homo sapi
c 4	2285	37.0	2800 93	AF04971 Homo sapi AF04971 Homo sapi
c 5	1035	16.8	1086 93	Z97171 Homo sapi Z97171 Homo sapi
c 6	978	15.9	1228 85	AB006686 Homo sapi AB006686 Homo sapi
c 7	647	10.5	1934 91	D88214 Homo sapi D88214 Homo sapi
c 8	640	10.4	1871 9	AX004457 Sequence AX004457 Sequence

	9	540	10.4	1871	9	AX004474	
10	640	10.4	1999	9	AR030962	Ax004474 sequence	
11	640	10.4	1999	9	AR056022	Sequence	
12	640	10.4	1999	9	AR059089	mapping group. Further information can be found at	
13	640	10.4	2000	97	HSU85257	http://www.sanger.ac.uk/HGP/Chr1/	
14	640	10.4	2002	9	A84847	This sequence has been finished according to sequence map criteria	
15	626	10.1	2061	93	HSU001620	as follows. An attempt is made to resolve all sequencing problems,	
16	604	9.8	1512	9	A84848	such as compressions and repeats, but not necessarily within known	
17	604	9.8	1512	9	AR030963	annotated human repeat sequence elements (e.g. Alu). Where the	
18	604	9.8	1512	9	AR066023	sequence 1 is ambiguous, there is an annotation using the 'unsure'	
19	604	9.8	1512	9	AR069090	feature key.	
20	594	9.6	2000	9	A84850	The true left end of clone 454G6 is at 1 in this sequence. The true	
21	558	9.0	1512	9	A84851	left end of clone 560B9 is at 79273.	
22	366	5.9	1969	10	AR020502	454G6 is from the library RP13 constructed at the Roswell Park	
23	330	5.3	1491	9	AR020503	Cancer Institute by the group of Pieter de Jong	
24	330	5.3	1491	10	I36470	For further details see http://bacpac.med.buffalo.edu/ .	
25	330	5.3	1491	10	Sequence 2	Location/Qualifiers	
26	189	3.1	228	88	AF007564	/r	
27	74	1.2	49868	78	AC090368	FEATURES	
28	74	1.2	73431	82	AP001857	source	
29	74	1.2	10290	84	HS838L14		
30	74	1.2	157454	82	AC001462		
31	74	1.2	166269	82	AP001187		
32	74	1.2	171980	82	AP000928		
33	74	1.2	176653	82	AP001857		
34	74	1.2	178100	82	AP001558		
35	74	1.2	203300	85	AC000134		
36	72	1.2	204022	73	AC068553		
37	70	1.1	153831	75	AC07168		
38	70	1.1	19927	86	AC007055		
39	65	1.1	11107	86	AC005924		
40	65	1.1	164245	68	AC04251		
41	65	1.1	175326	83	CNS01DQ		
42	64	1.0	136886	89	ALI38889		
43	64	1.0	168210	87	AC018719		
44	64	1.0	168544	89	AC018719		
45	64	1.0	179040	67	AC022732		
						Homo sapiens	
						/organism="Homo sapiens"	
						/db_xref="taxon: 9606"	
						/chromosome="1"	
						/map="1q24"	
						/clone="RPC3-454G6"	
						/clone_11b="RPC1-3"	
						435..472	
						/note="19 copies of 2 mer 82 & conserved"	
						1914..>368	
						/note="match: multiple ESTs"	
						match: R56676 AA043668 W63639 F12081 AA046699	
						match: A02952 AA043668 W63634 R36066	
						match: AA13561 F02025 AA131540 W00634 R36066	
						match: AA13383 AA163561 AA9173 AA174814 AA057059	
						match: AA129084 W47062 AA043955 AA341783 AA353681	
						match: AA046487 AA369741 H08313 AA186895 H32730	
						match: H0333 H08236 N42052 D61944 R27102 N32353	
						match: N30491 AA307150 AA192"	
						3703..3746	
						/note="22 copies of 2 mer 89 & conserved"	
						4051..4183	
						/note="Alusq repeat: matches 1..133 of consensus"	
						repeat_region	
						4200..4502	
						/note="Alusq repeat: matches 2..301 of consensus"	
						4659..4831	
						/note="Alusq repeat: matches 2..194 of consensus"	
						repeat_region	
						5216..5335	
						/note="Alusq repeat: matches 132..1 of consensus"	
						incomplete repeat"	
						repeat_region	
						7759..7907	
						/note="MTR repeat: matches 174..1 of consensus"	
						7933..9328	
						/note="TIGER1 repeat: matches 1..1472 of consensus"	
						9332..9656	
						/note="Alusq repeat: matches 1..289 of consensus"	
						9639..10335	
						/note="Alusq repeat: matches 1469..2174 of consensus"	
						10343..110642	
						/note="Alusq repeat: matches 1..300 of consensus"	
						10643..10856	
						/note="TIGER1 repeat: matches 2175..2417 of consensus"	
						<10903..10985	
						/note="endogenous retroviral sequence"	
						10949..11384	
						/note="LRR2 repeat: matches 31..449 of consensus"	
						11583..15870	
						/note="18 copies of 2 mer 83 & conserved"	
						16856..17286	
						/note="LRR2 repeat: matches 31..449 of consensus"	
						17287..17574	
						/note="Alusq repeat: matches 15..300 of consensus"	
						18294..18650	
						/note="TIR1B repeat: matches 358..1 of consensus"	
						18877..19180	
						repeat_region	

This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre chromosome 1 mapping group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr1/>. This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence 1 is ambiguous, there is an annotation using the 'unsure' feature key.

The true left end of clone 454G6 is at 1 in this sequence. The true left end of clone 560B9 is at 79273.

454G6 is from the library RP13 constructed at the Roswell Park Cancer Institute by the group of Pieter de Jong

For further details see <http://bacpac.med.buffalo.edu/>.

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variations annotated may not be found in the sequence submission corresponding to the overlapping clone as we submit sequences with only a small overlap as described above.

REFERENCE HS4546/c
LOCUS HS4546
DEFINITION Human DNA sequence from PAC 454G6 on chromosome 1q24. Contains trabecular meshwork inducible glucocorticoid response protein, TIGR, myocilin, ESTs and STS.
ACCESSION AC005924
VERSION 298750
KEYWORDS
SOURCE JOURNAL
ORGANISM Homo sapiens

Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo. REFERENCE 1
AUTORS Deadman, R.
TITLE Direct Submission
SUBMITTED (27-OCT-1997) Chromosome 1 Project Group
VERSION 298750.1 GI:2887277
KEYWORDS 1q24; myocilin; TIGR.
SOURCE human.
ORGANISM Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
COMMENT On Feb 14, 1998 this sequence version replaced gi:2455060.
IMPORTANT: This sequence is not the entire insert of clone 454G6. It may be shorter because we only sequence overlapping sections once or longer because we arrange for a small overlap between neighbouring submissions.

Db	4801	TITLE	Zody, M.
Oy	4861	JOURNAL	Direct submission
Db	4861	COMMENT	Submitted (28-FEB-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
Oy	4921	On May 25, 2000 this sequence version replaced g1.7249345. All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997)	
Db	4921	http://ftp.genome.washington.edu/RM/RepeatMasker.html	
Oy	4981	Center: Whitehead Institute/MIT Center for Genome Research	
Db	4981	Center code: WIBR	
Oy	5041	Web site: http://www.seq.wi.mit.edu	
Db	5041	Contact: sequence_submissions@genome.wi.mit.edu	
Oy	5041	----- Project Information	
Db	5041	Center project name: L7153	
Oy	5101	Center clone name: 138.F.3	
Db	5101	----- Summary Statistics	
Oy	5161	Sequencing vector: M13; M7815; 100% of reads	
Db	5101	Chemistry: Dye-terminator Big Dye; 100% of reads	
Oy	5221	Assembly program: Phrap; version 0.960731	
Db	5221	Consensus quality: 151483 bases at least Q40	
Oy	5281	Consensus quality: 160439 bases at least Q30	
Db	5281	Consensus quality: 164708 bases at least Q20	
* Insect size: 157000; agarose-fp			
* Insert size: 167725; sum-of-contigs			
* Quality coverage: 4.1 in 020 bases; agarose-fp			
* Quality coverage: 4.1 in 020 bases; sum-of-contigs			
* NOTE: This is a 'working draft' sequence. It currently consists of 28 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.			
* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.			
* * 1 1300: contig of 1300 bp in length			
* * 1301 1400: gap of 100 bp			
* * 1401 2412: contig of 1012 bp in length			
* * 2413 2512: gap of 100 bp			
* * 2513 4147: contig of 1635 bp in length			
* * 4148 4247: gap of 100 bp			
* * 4248 5455: contig of 1208 bp in length			
* * 5456 5555: gap of 100 bp			
* * 5556 6997: contig of 1442 bp in length			
* * 6998 7097: gap of 100 bp			
* * 7098 8534: contig of 1437 bp in length			
* * 8535 8634: gap of 100 bp			
* * 8635 10164: contig of 1530 bp in length			
* * 10165 10264: gap of 100 bp			
* * 10265 10487: contig of 223 bp in length			
* * 10488 10587: gap of 100 bp			
* * 10588 12420: contig of 1833 bp in length			
* * 12421 12520: gap of 100 bp			
* * 12521 14843: contig of 2323 bp in length			
* * 14844 14943: gap of 100 bp			
* * 14944 17265: contig of 2322 bp in length			
* * 17266 17365: gap of 100 bp			
* * 17366 20071: contig of 2706 bp in length			
* * 20172 20171: gap of 100 bp			
* * 20959 29988: contig of 100 bp			
* * 23532 23631: gap of 100 bp			
* * 23632 26137: contig of 2505 bp in length			
* * 26138 26237: gap of 100 bp			
* * 26238 29858: contig of 3621 bp in length			
* * 29959 36225: contig of 6267 bp in length			
* * 36226 36325: gap of 100 bp			
* * 36326 44015: contig of 7690 bp in length			
* * 44016 44115: gap of 100 bp			
* * 44116 44943: contig of 5378 bp in length			
* * 44954 45933: gap of 100 bp			
* * 56796: contig of 7203 bp in length			

REFERENCE	Nobuyoshi Shimizu, Keio University School of Medicine, Department of Molecular Biology; 35 Shinanomachi, Shinjuku, Tokyo 160, Japan (E-mail:shimizu@med.med.keio.ac.jp), Tel:03-3351-2370, Fax:03-3351-2370 2 (sites)
AUTHORS	Kubota, R., Noda, S., Wang, Y., Minoshima, S., Asakawa, S., Kudo, J., Mashima, Y., Oguchi, Y. and Shimizu, N.
TITLE	A novel myosin-like protein (myocilin) expressed in the connecting cilium of the photoreceptor; molecular cloning, tissue expression, and chromosomal mapping
JOURNAL	Genomics 41 (3), 360-369 (1997)
MEDLINE	97312692
COMMENT	On Nov 20, 1997 this sequence version replaced 91:2104213. Sequence updated (17-Nov-1997).
FEATURES	Location/Qualifiers
source	1..1934
CDS	<pre> </product>"myocilin" </protein>_id="BA22531_1" </db_xref>"GI:2627177" </translation>"MRFCCARCSFGPEMPAVOLLACLWWDGARTAQLRKANDOS GRCYTFSAVSPRESSQPSQMSVILNQDSTDQRDLERKARSLLESILHLQ TLDQAAKRPQETQELQRELTGRERDQETQTELETYSNLURKSYLEEKRLR QENENLARLKESSQEVALLRRSCOPOTDTARAVPESREVSNWLQDFOLKE LTPVANASRILKESSQEVALLRRSCOPOTDTARAVPESREVSNWLQDFOLKE TYPTQETWTWIDYVGTDQRQFVEYDLSIQFMQYPSKHLPRPLESGAVVWSGL YFQCAESRVRVILENTENKEAPEIPIGAGHOFPSYGGYDIDLADEAGLWVY STBEAKGAIVLUSKLNPENELQWTKRQVANAFATTCIGCTTUVTSADATVN FADTGTGSKTLTIPFKRYKISSMDINPLERKFALWDNLNAVYDQDKSM" 104..105 </citation>[2] </replace>"gaga" 317..361 </note>"leucine zipper-like motif 1" 413..562 </note>"leucine zipper-like motif 2" 178..1783 1928..1933 </note>"leucine zipper-like motif 2" 510..503 503..525 525..396 t..t </pre>
BASE COUNT	
ORIGIN	
Query Match	10.5%; Score 647; DB 91; Length 1934;
Best Local Similarity	100.0%; Pred. No. 0; Mismatches
Matches	647; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy	5294 gcaatgcacgtttccaaaggaaagctccaaacabcccttcgtcaatgggtttcttgtc 5353
Db	22 GCACAGCACAGCTTCCAGAGGAAGCCTCCAAAGCCTCGCATGAGGTCTCTGTC 81
Oy	5354 acgttgctcagaatggcgttagatgcacgtttccacgtgtcgatgtttggctgc 5413
Db	82 ACGRGCTGAGCTTGGCCTGAGATGCCAGCTCCACGCTGCCTCTGGCTGCCR 141
Oy	5414 gggtgtggatgtggggccggacgtcgactcgaaaggccatgaccaggatggccc 5473
Db	142 GGTTGGGATGTTGGGGCCAGCTCGCTCAGGAACGCCAATGACCAAGGGGCC 201
Oy	5474 atgcagatatacttcgttgccatccaaatccatcgacgtggccatcgccggggcca 5533
Db	202 ATGCCAGTATACCTTCAGTGGCAGTCCAAATGATCCAGTCGCCAACAGAGCCA 261
Oy	5534 ggccatgtcgatccatcttcacagatggacggacggacccaaacgttagactgtgg 5593
Db	262 GGCCATGTCATCCATTAACCTACAGAGAGACGCCAACGCTTAGACTGGGA 321
RESULT	8
AX00457	
LOCUS	AX004457
DEFINITION	Sequence 1 from Patent WO9916898.
ACCESSION	AX004457
VERSION	AX004457.1
PAT	24-AUG-2000
KEYWORDS	
ORGANISM	Homo sapiens
REFERENCE	1 (bases 1 to 1871)
AUTHORS	Ancill, J.L. and Cote, G.
TITLE	Molecular diagnostic of glaucomas associated with chromosomes 1, and method of treatment thereof
JOURNAL	Patent: WO 9916898 A 1 08-APR-1999; ANCILL JEAN LOUIS (CA); COTE GILLES (CA)
FEATURES	Location/Qualifiers
source	1..1871 </note>"Homo sapiens" </db_xref>"taxon:9606"
BASE COUNT	499 a..483 c..507 g..382 t..
ORIGIN	
Query Match	10.4%; Score 640; DB 9; Length 1871;
Best Local Similarity	100.0%; Pred. No. 0; Mismatches
Matches	640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy	5301 agacgtttcccgaggagccatccaaacccggccatggatgggtttctgtggcacccatggc 5360
Db	1 AGACGTTCAGAGGAAGCCTTCAACAGCCATCTGCAATGAGGTCTCTCTGTC 60
Oy	5361 tgcacgtttggcgttagatgcacgtgtccacgtgtcgatgtttgggtttgg 5420
Db	61 TCCAGCAGTTGGCCTGAGATGCCAGTGTCAGCTGCTGCTGCTGGCTGGCTGG 120
Oy	5421 gatgtgtggggccaggacgtcgactcgatggggccatgaccaggatggccatccag 5480
Db	121 GATGTCGGGGCAGGACAGCTCAGTCAGCTGAGGCCAATGACCAAGGGCTGCCAG 180
Oy	5481 tataccatgtgtccatcgccatgtgaatccatcgatgtggccatcgccggcatgtgg 5540
Db	181 TATACCTTCAGITGGCCAGTCCAAATGATCCAGTGGCCAGAGGAGGCCATG 240
Oy	5541 tcgtatccataacttcacatggacggacggccacccaaaccttaaacctcgatggcacc 5600
Db	56..561

Search completed: November 9, 2001, 01:57:17
Job time: 46361 sec